References are Secondary to Appropriate Normalization

This is true for any technology

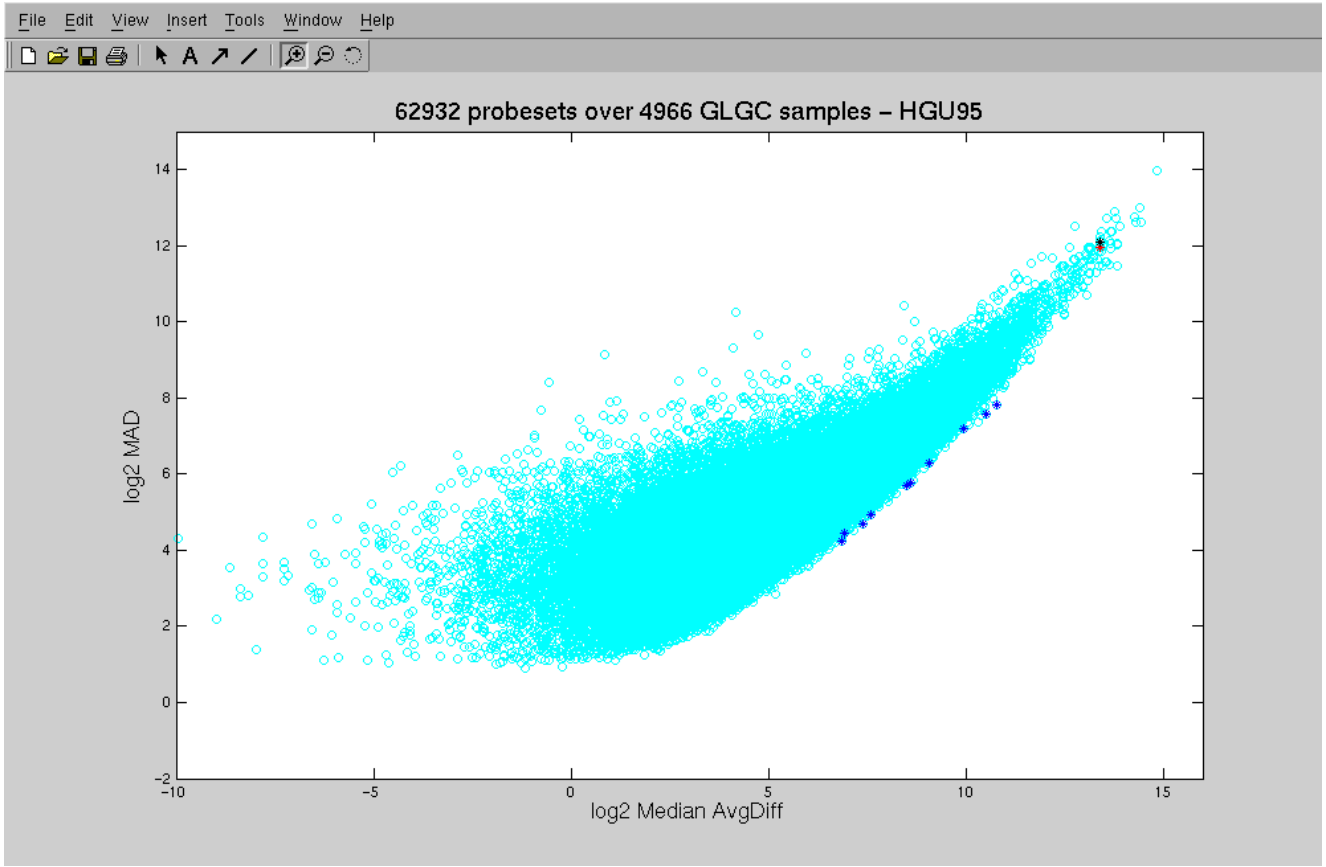
- Appropriate normalization of data is important for accurate sample comparison
- Normalization should account for any differences in sample
  - mass or
  - labeling/detection efficiency

Normalization for Real-Time qPCR

(a bit more tricky than for high density microarrays)

- Normalization is necessary to compare a gene of interest across different samples
- The goals of normalization are to control for variation:
  - Total mass of RNA analyzed
    (requires analysis of endogenous gene)
  - Efficiency of Amplification
    (at 95% efficiency only ~ 47% of product is generated)
- Normalization in Real-Time PCR is generally achieved by ratio-ing the gene of interest to a single or small number of "verified invariant gene(s)"
- Selection of a "verified invariant gene" is important for data analysis

Selection of “Least” Variant Genes from GeneLogic® Affymetrix® Dataset

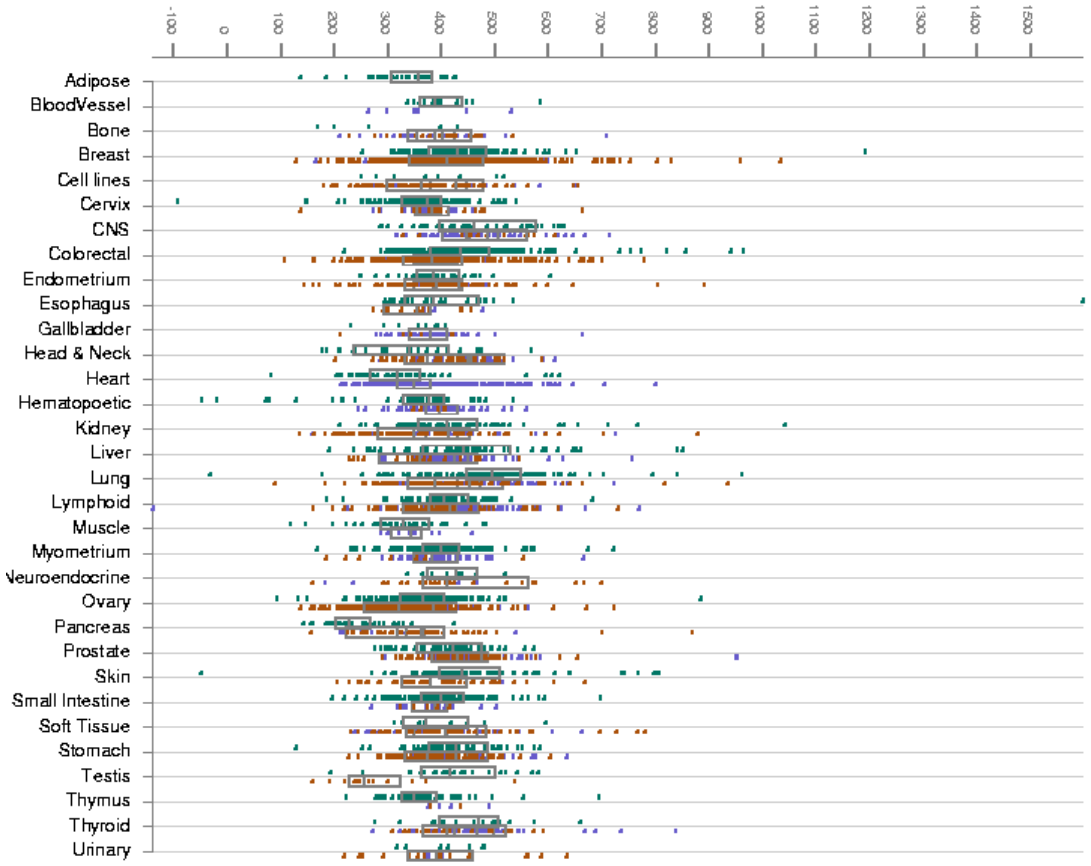


Array Expression Variability

Gene Id	Mean	Median	StdDev	%CV
Beta Actin	3016	2433	2189	**72**
GAPDH	2884	2354	2003	**69**
A	200	189	74	37
H	131	125	43	33
F	550	530	168	31
D	122	118	37	30
G	1027	988	307	30
E	402	390	109	27
I	174	168	46	27
B	372	365	95	26
J	1544	1494	370	24
C	1792	1738	402	22

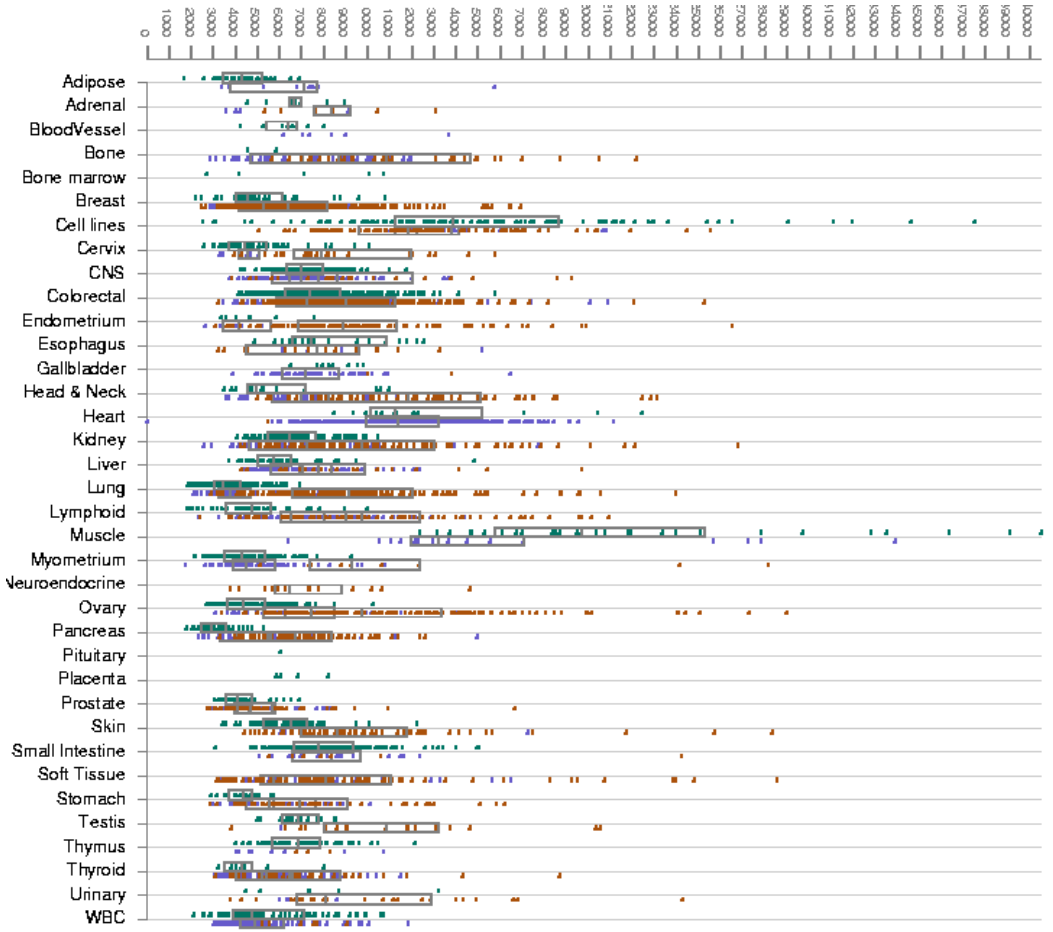
Genelogic® data: 32 tissue types with ~5000 samples including disease

Gene E

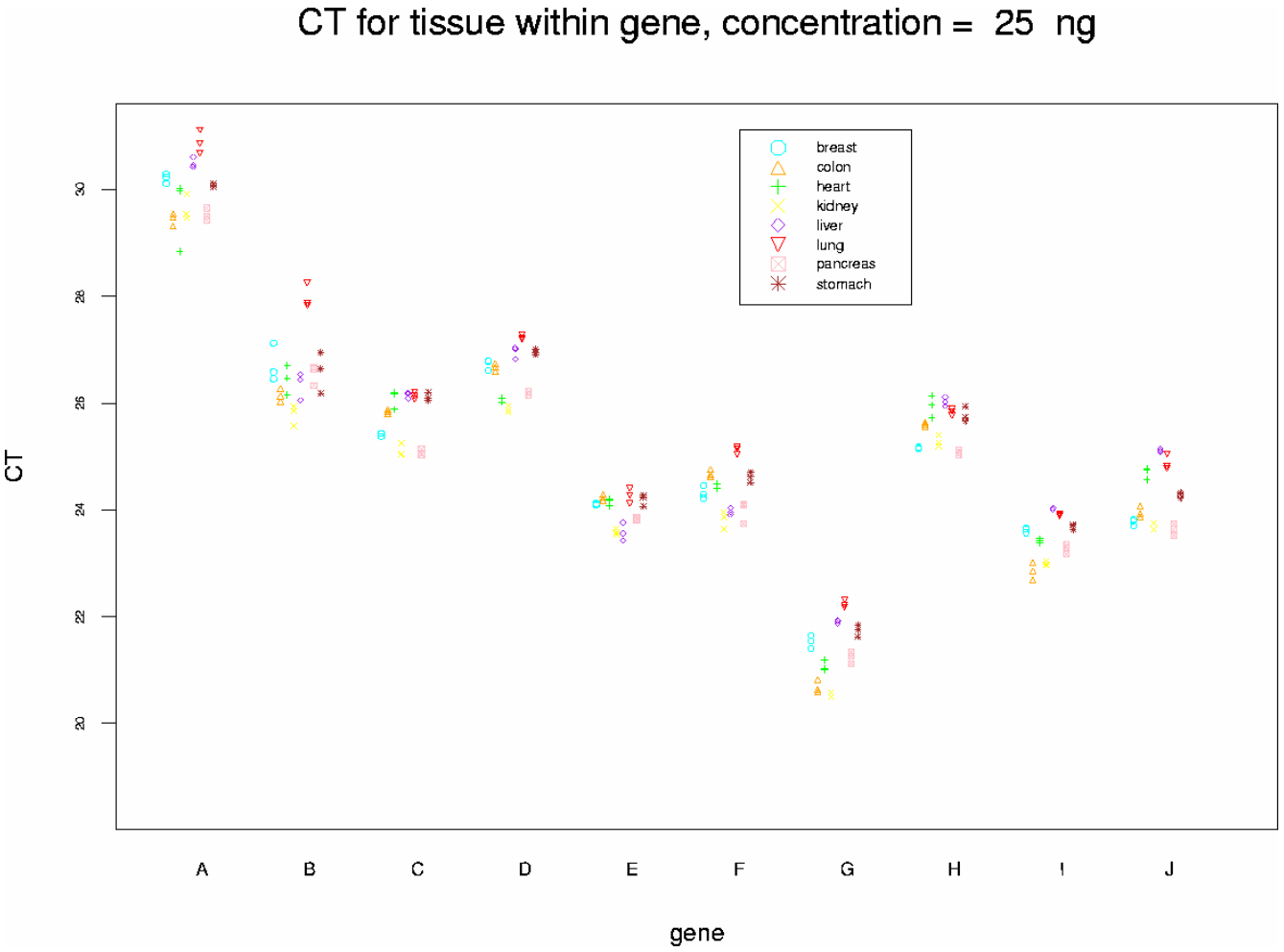


35325_at gen.NM_016322 NM_016322 Homo sapiens RAB14, member RAS oncogene family [RAB14], mRNA. 4106 bp, mRNA, linear, PRI 04-APR-2002

GAPDH


Adipose
Adrenal
BloodVessel
Bone
Bone marrow
Breast
Cell lines
Cervix
CNS
Colorectal
Endometrium
Esophagus
Gallbladder
Head & Neck
Heart
Kidney
Liver
Lung
Lymphoid
Muscle
Myometrium
Neuroendocrine
Ovary
Pancreas
Pituitary
Placenta
Prostate
Skin
Small Intestine
Soft Tissue
Stomach
Testis
Thymus
Thyroid
Urinary
WBC
AFFX-HUMGAPDH/M33197_3_at M33197 Human glyceraldehyde-3-phosphate dehydrogenase [GAPDH] mRNA, complete cds [_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively]

Variation of 10 Selected Genes Across Multiple Normal Tissues



Invariant Expression of Gene E in Treated and Control Samples

	E		
	200 ng	**50**	**12.5**
105B	19.75	21.05	23.34
141A	19.11	21.2	22.98
148B	19.56	21.2	23.25
161A	19.73	21.36	23.44
169B	19.31	21.17	23.25
Average	19.49	21.20	23.25
Sdev	0.277	0.111	0.171
Max	19.75	21.36	23.44
Min	19.11	21.05	22.98
Max-Min	0.64	0.31	0.46

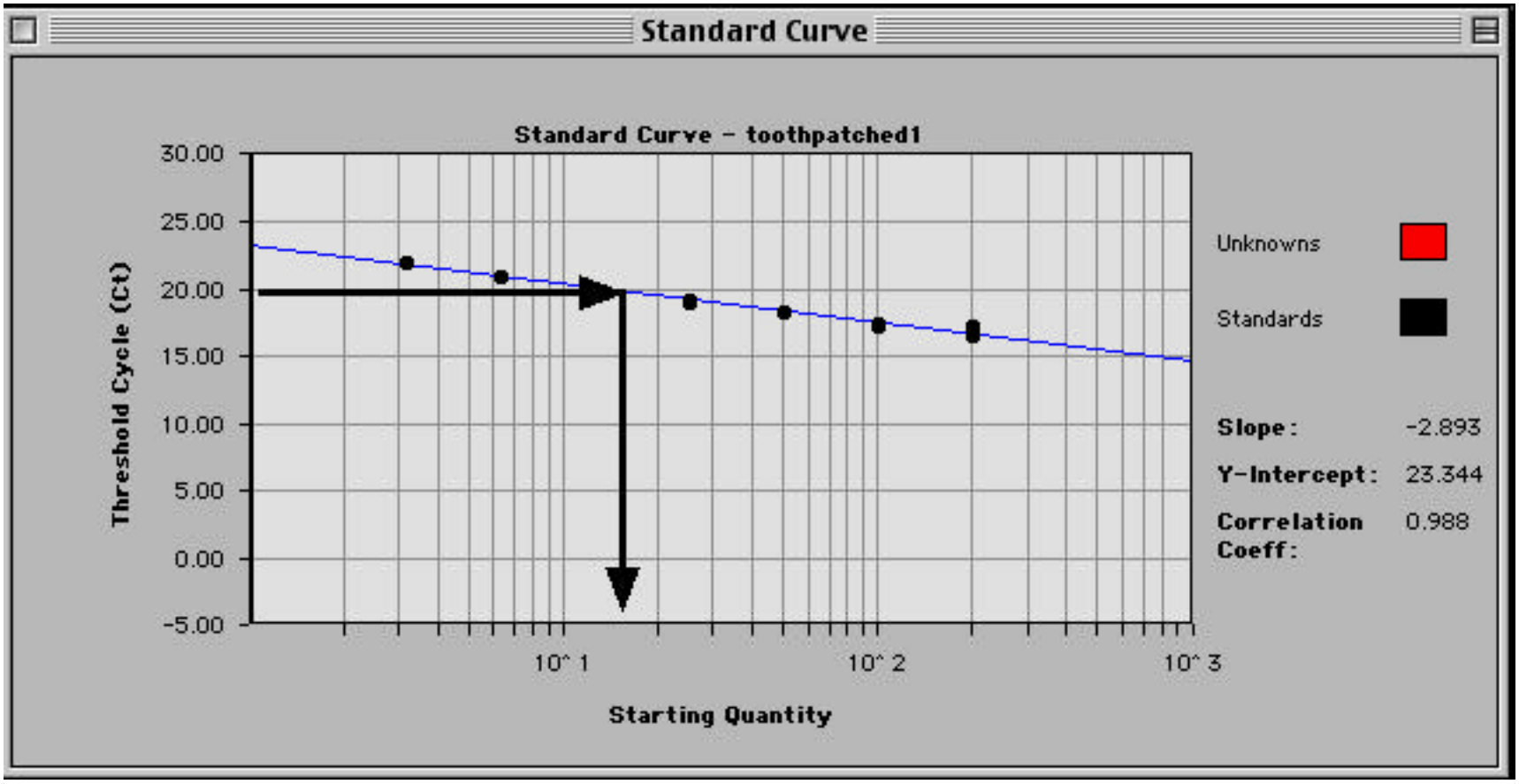
A= untreated tumor
B= chemo treated tumor

Quantitative Analysis Using Real-time qPCR and a Reference Sample

- We prefer to use a reference sample to generate a standard curve (usually a cell or tissue that expresses the genes of interest, this sample usually varies with projects)
  - A total genome universal reference would be “most wonderful”
  - This would permit direct comparison of different experiments
  - This would permit different primers and probes for the same gene to be directly compared
- The genes of interest are analyzed in this reference standard curve
  - The standard curve is used to generate a linear equation which is used to determine quantitative amount of gene in sample of interest “relative” to standard
  - The standard curve also demonstrates linear dynamic range of the assay
- Genes of interest are then ratio’d to a verified invariant normalization gene for relative comparison across samples

Use Ct of Sample of Interest

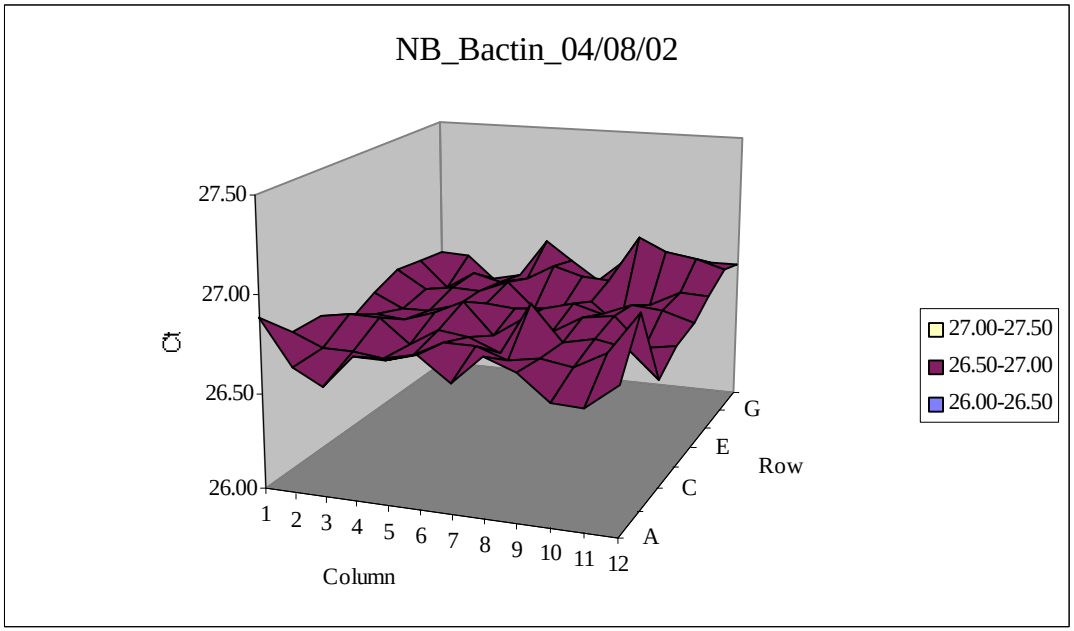
solve for relative quantity comparing to reference standard



Reference Standard is usually an RNA from a cell line or tissue expressing genes of interest

Real-time PCR Instrument Performance Validation

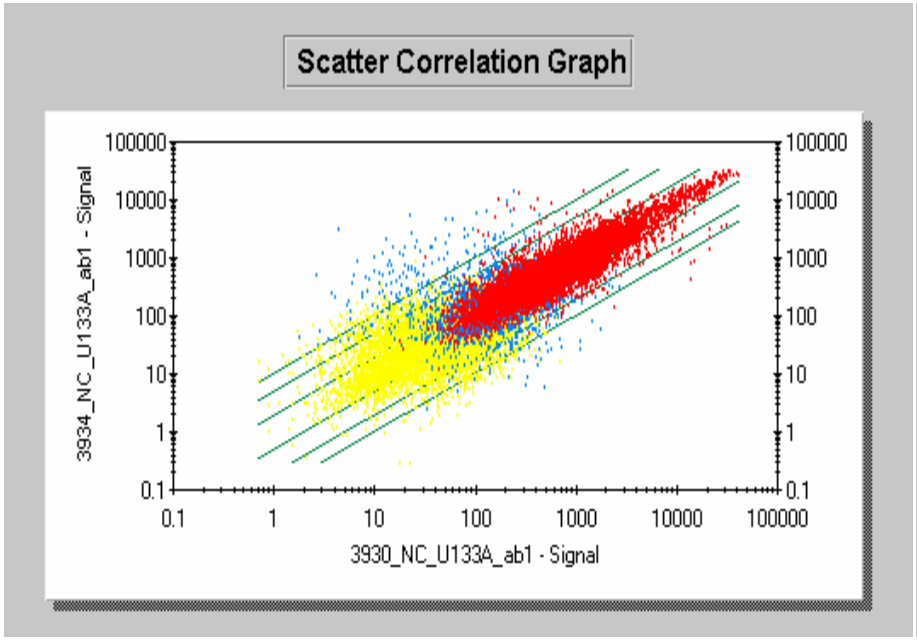
Analysis of 96 well replicate of Beta-actin

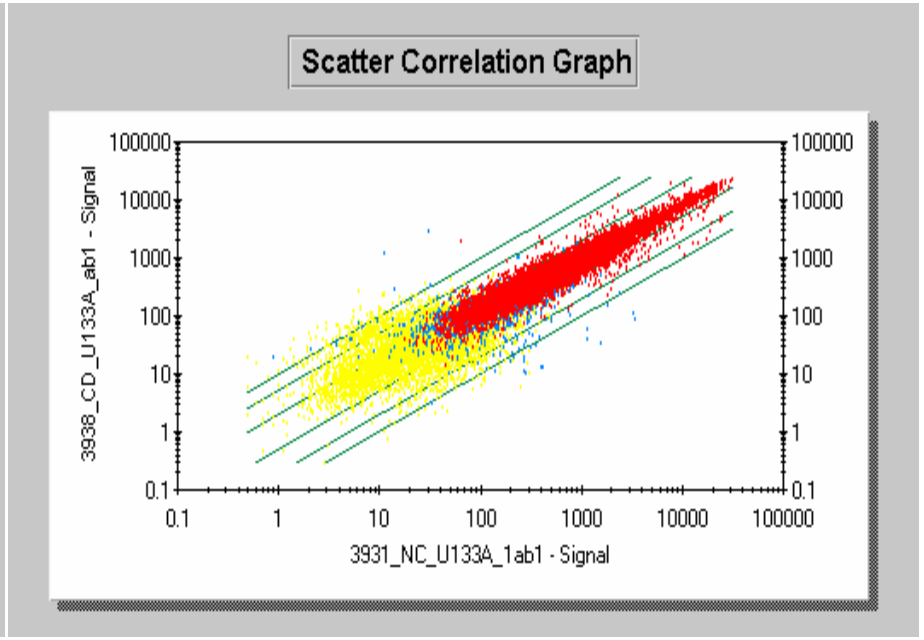


Spec: Ct min to max < 1
stdv less than 0.15

Microarray Normalization

Most Genes Don't Change, therefore Global Scaling is an adequate choice





Normal Colon #1 vs. Normal Colon #2

Normal Colon vs. Crohn's Disease Colon

Affymetrix® Technology Performance Validation

- Use of Spikes
  - Pre-label monitors labeling enzymology
  - Post-label monitors hybridization and instrument performance
- 5’and 3’ probes help monitor sample integrity
- It would be great to have a reference chip for instrument calibration

Cross Platform Comparisons

- Is it necessary to have exact quantitative comparisons across platforms?
  - Maybe not for research discovery within an organization
    - But it would be nice to compare outside data
  - Most definitely for diagnostic (e.g. viral burden, single gene assays)
  - Maybe or maybe not for class prediction

TaqMan® vs. Affymetrix® MW-Data (>95% TaqMan® Verification*)

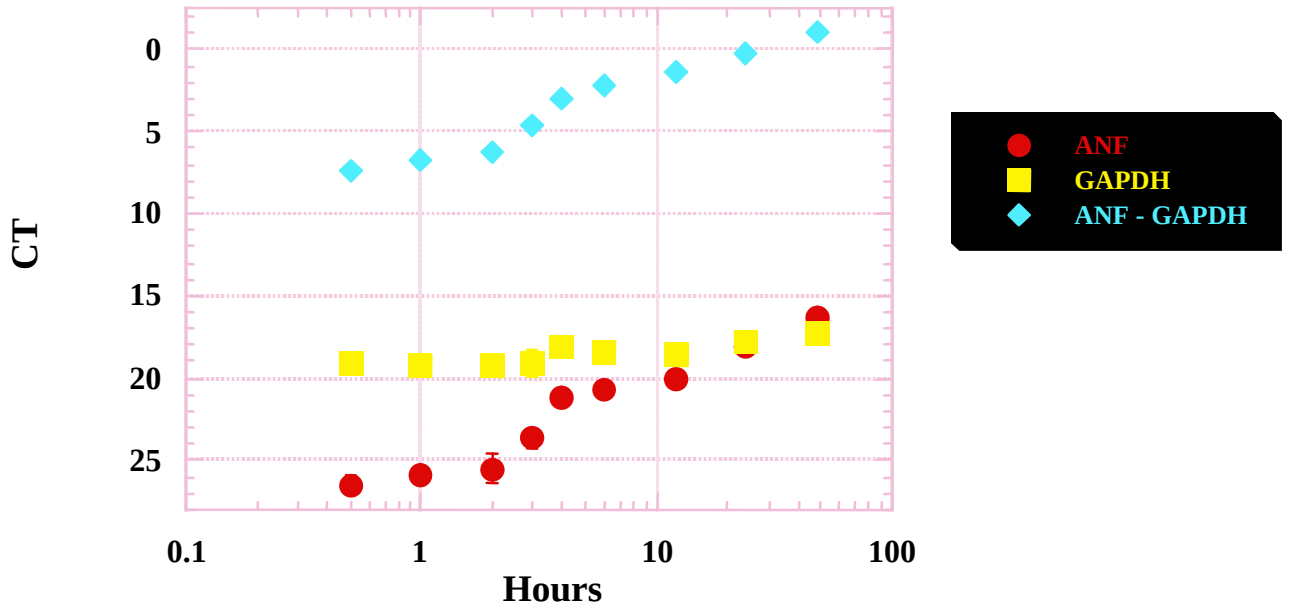
Function	Gene	Array	TaqMan
		MI/Sham	MI/Sham
Natriuretic Factor	ANF	2.5	5.4
Extracellular Matrix	Collagen I	3.3	3.4
	Collagen III	2.3	3.4
	lysyl oxidase	2.5	3.4
	"TIMP1"	4.3	1.8 p=0.06
Inflamation/Wound Healing	TGFb-3	2.5	2.1
	VCAM-1	4.4	2.4
	LPS BP	5.9	10.2
IGF-1 Axis	IGF-1	4	2.4
Muscle Proteins	SM22	(27.7) (3.5)	2.3
	"acidic calponin"	2.3	1.2
Ion Binding/Channels	Isk	7.5	2.4 p=0.07
	p9Ka	2.7	2.5
Growth Factors/Inhibitors	TRH	8.7	2.9
	tumor-sup gene	2.7	3
Steroid Binding/Metabolism	P450 1B1	10.3 (7.4)	3.7
Others	ost.-sp. factor-2	4.6	2.9
	unknown est	1.8	1.5

* To date virtually all genes (>95%) found by MW confirm

Hurdles of Quantitation

- Selection and verification of invariant genes for real-time qPCR
- Sample quality
  - does every gene degrade with equal assay results?
- Complex Tissues
  - reporting gene expression per mass of tissue, per cell, per cell type within tissue

PGF2α Time Course of Induction of ANF mRNA



GAPDH Values of Control and Hypertrophied Myocytes

Is GAPDH Expression Changing or
Is RNA Mass Changing or
Are the Number of Cells Changing

Control Cells	CT	Mean	STDEV
A	17.85		
B	18.26	17.94	0.257
C	17.65		

PGF2a Cells	CT	Mean	STDEV
D	16.63		
E	16.54	16.64	0.135
F	16.55		

GAPDH is Invariant with RNA Mass

But Increases Expression upon Hypertrophy

Cell Number	Average Cells	Total RNA	Average RNA	CT/15ng	Average CT
Control Cells					
3.6 x 10^6		18.01 ug		16.25	
3.3 x 10^6	3.5 x 10^6	14.05	16.08	16.30	16.28
3.6 x 10^6		16.18		16.29	
PGF2a					
3.8 x 10^6		27.34 ug		16.47	
4.1 x 10^6	3.7 x 10^6	32.16	30.92	16.54	16.55
3.2 x 10^6		33.27		16.65	